OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen, Ltd.

GenCore version 5.1.6

Run on: August 13, 2003, 04:28:29 ; Search time 3057.14 seconds
(without alignments)
11440.147 Million cell updates//sec

Title: US-09-818-143-6
Perfect score: 1439
Sequence: 1 gagtatccctgtttaatca.........agaaaaaaaaaaaaaaaa 1439

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215223B056 residues

1 number of hits satisfying chosen parameters: 4556278A

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
```
          EST:*
      1:  em_estba:*
      2:  em_esthum:*
      3:  em_estin:*
      4:  em_estmu:*
      5:  em_estov:*
      6:  em_estpl:*
      7:  em_estro:*
      8:  em_htc:*
      9:  em_htc1:*
     10:  gb_est1:*
     11:  gb_est2:*
     12:  gb_est3:*
     13:  gb_est4:*
     14:  gb_est5:*
     15:  em_estom:*
     16:  em_estfun:*
     17:  em_estom:*
     18:  em_gss_hum:*
     19:  em_gss_inv:*
     20:  em_gss_pln:*
     21:  em_gss_vrt:*
     22:  em_gss_fun:*
     23:  em_gss_mam:*
     24:  em_gss_mus:*
     25:  em_gss_pro:*
     26:  em_gss_rod:*
     27:  em_gss_phg:*
     28:  em_gss_vrl:*
     29:  gb_gss1:*
         gb_gss2:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1439	100.0	1439	10	AW888220	AW888220 MXRA3 Hum
2	1292	89.8	2310	11	BC032745	BC032745 Homo sapi
c 3	813.6	56.5	922	9	AL521507	AL521507 AL521507
c 4	804.6	55.9	922	9	AL521506	AL521506 AL521506

RESULT 1
AW888220

LOCUS   AW888220   1439 bp   mRNA   linear   EST 05-MAR-2001
DEFINITION AW888220 MXRA3 Human matrix tissue expression library Homo sapiens cDNA
            clone Incyte 1556751 similar to MXRA3 Matrix remodeling associated
            gene 3, mRNA sequence.
ACCESSION   AW888220
VERSION   AW888220.1 GI:8050231
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1439)
AUTHORS   Walker,M.G. and Volkmuth,W.
TITLE   Matrix-remodeling associated genes identified by co-expression
JOURNAL   Unpublished
COMMENT   Contact: Michael Walker
            Incyte Genomics
            3160 Porter Drive, Palo Alto, CA 94304, USA
            Tel: 650 845-5771
            Fax: 650 621 8514
            Email: mwalker@incyte.com.
FEATURES   Location/Qualifiers
     source   1..1439
                     /organism="Homo sapiens"
                     /mol_type="mRNA"

ALIGNMENTS

```
c  5  735.2  51.1  741  14  CA430710  CA430710  UI-H-FL1-
c  6  706    49.1  738  10  BG748266  BG748266  602705508
c  7  701.4  48.7  703  12  BO009860  BO009860  UI-H-ED1-
c  8  699.8  48.6  703  12  BM980547  BM980547  UI-CF-EN1
c  9  698.6  48.5  668  12  CB851585  CB851585  UI-CF-EN1
c 10  699.6  48.2  731  14  BX325388  BX325388  BX325388
c 11  687    47.7  703  10  BG678659  BG678659  602624453
c 12  677.4  47.1  733  12  BM974268  BM974268  UI-CF-EC1
c 13  635.2  44.1  674  14  CA440929  CA440929  UI-H-ED0-
c 14  633.2  43.5  668  12  BM701957  BM701957  UI-E-CQ1-
c 15  624.8  43.4  673  13  BQ448042  BQ448042  UI-H-EU1-
c 16  624.8  43.4  683  13  BU627928  BU627928  UI-H-FG0-
c 17  607    42.2  760  13  BU170946  BU170946  AGENCOURT
c 18  603.6  41.9  766  12  BM995562  BM995562  UI-H-DT1-
c 19  596    41.4  766  13  CD104758  CD104758  AGENCOURT
c 20  590.4  41.0  636  13  BU729623  BU729623  UI-E-CO1-
c 21  582.6  40.5  602  13  BU683800  BU683800  UI-CF-EN0
c 22  574.4  39.9  594  14  CB052399  CB052399  NISC_gl07
c 23  560.2  38.9  628   9  AA115295  AA115295  2108e72.r
c 24  555.6  38.6  605  14  CA777565  CA777565  1p18a07.y
c 25  543.6  37.8  605  14  CA777311  CA777311  1p18a07.x
c 26  541    37.6  587  14  CB306370  CB306370  UI-CF-EN1
c 27  537.6  37.4  571  12  BM710342  BM710342  UI-CF-EN1
c 28  537    37.3  571  12  BM090782  BM090782  ig15a08.x
c 29  534.6  37.2  542  12  BQ019900  BQ019900  UI-s-CQ1-
c 30  532    37.0  540  12  BM090553  BM090553  ig15a08.y
c 31  532    36.9  535   9  AL049024  AL049024  DKFZp434O
c 32  527    36.6  535  10  BF724489  BF724489  bx05e03.y
c 33  527    36.6  538   9  BM671762  BM671762  UI-E-CQ1-
c 34  525.8  36.5  642  12  BU633007  BU633007  UI-H-DF0-
c 35  522    36.3  530   9  AA625435  AA625435  af72a07.r
c 36  517.4  36.2  508  14  CA951879  CA951879  1q15b12.x
c 37  507    35.2  558   9  AI983849  AI983849  wt51e05.x
c 38  504    35.0  649  10  BE966909  BE966909  60160928
c 39  499.8  34.7  570  10  AW963887  AW963887  EST375960
c 40  496.6  34.5  501   9  AA780019  AA780019  zj88g05.s
c 41  496.2  34.5  621  14  W84522    W84522    zd90f10.s1
c 42  494    34.3  516   9  AA873315  AA873315  oh76h01.s
c 43  483.2  33.6  518  12  BM984490  BM984490  UI-CF-DU1
c 44  479.2  33.3  501   9  AW069314  AW069314  cr44b01.x
c 45  476.6  33.1  479   9  AI926527  AI926527  wo46a11.x
```

```
BASE COUNT      392 a    336 c    338 g    373 t
ORIGIN
                /db_xref="taxon:9606"
                /clone="Incyte 1556751"
                /clone_lib="Human matrix tissue expression library"

Query Match          100.0%;   Score 1439;  DB 10;  Length 1439;
Best Local Similarity 100.0%;  Pred. No. 4e-208;
Matches 1439; Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy       1 GAGTATCCCTTGTTAATCACTTTGTGGTAAAAGAGACCTTGGGTCAGTCTGCCTCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GAGTATCCCTTGTTAATCACTTTGTGGTAAAAGAGACCTTGGGTCAGTCTGCCTCA 60

Qy      61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTCACTCTATTTCTTCTCTGTCTCAAGAA 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTCACTCTATTTCTTCTCTGTCTCAAGAA 120

Qy     121 AGAAGAAAAAAAGAGAGACAAAATTACCCAGAAAACCCCTCCCCCACAATGGAGGCCTTGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGAAGAAAAAAAGAGAGACAAAATTACCCAGAAAACCCCTCCCCCACAATGGAGGCCTTGG 180

Qy     181 CAAATGTTAATTTCCTTAGAAAAATCCTTCAGACCTGAACCCTTCCCCTGTGAAAAACCAAAACAAAGGAATTCTGGGCT 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAAATGTTAATTTCCTTAGAAAAATCCTTCAGACCTGAACCCTTCCCCTGTGAAAAACCAAAACAAAGGAATTCTGGGCT 240

Qy     241 CTCAGGGTGGGTTCTTCGCGTCCCCGCCCCGGCCCCCCCAGACTATGTGTCACGGGCCGTCCT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTCAGGGTGGGTTCTTCGCGTCCCCGCCCCGGCCCCCCCAGACTATGTGTCACGGGCCGTCCT 300

Qy     301 GTTCCTTCCCCGGGACTCCAGAATTTCTCTCCTCCCAAAGGAAAGAAAACAGGGGCATGGCTT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GTTCCTTCCCCGGGACTCCAGAATTTCTCTCCTCCCAAAGGAAAGAAAACAGGGGCATGGCTT 360

Qy     361 GTTGGGCAAAACGCAGGGCCGGCCTCCCAAAAAACCCATGTGTGTACGATTAAAAGTTTGGGCC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTTGGGCAAAACGCAGGGCCGGCCTCCCAAAAAACCCATGTGTGTACGATTAAAAGTTTGGGCC 420

Qy     421 GTCCCCAGGGCCTCCCAGCGCCAAACTTAAAGAGACACAGGGGCTTTGCTGAAAAACCAAAACATGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTCCCCAGGGCCTCCCAGCGCCAAACTTAAAGAGACACAGGGGCTTTGCTGAAAAACCAAAACATGG 480

Qy     481 GCCAGCTGGGCGCTTTTTAACAACCTAGAGACTTTCCGGGACTGCCTGGGAACACAGAGCCTGCG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GCCAGCTGGGCGCTTTTTAACAACCTAGAGACTTTCCGGGACTGCCTGGGAACACAGAGCCTGCG 540

Qy     541 GGAAAACGGGGCTTTGCCAGAGACACACTCACAGTTTCTTCTTCAGGACAGTTTCCTTCATGGGCCTGTTTTGGTCCCCTAA 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGAAAACGGGGCTTTGCCAGAGACACACTCACAGTTTCTTCTTCAGGACAGTTTCCTTCATGGGCCTGTTTTGGTCCCCTAA 600

Qy     601 GAATCTCCACATCATTGTCTTCTTGTTTCTTGTTGTGCCTTTTCCTTGGTGAGCAACAGAAAGGGAAGG 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GAATCTCCACATCATTGTCTTCTTGTTTCTTGTTGTGCCTTTTCCTTGGTGAGCAACAGAAAGGGAAGG 660

Qy     661 GTTCCAAGCCTCTAAAAATGTGCTTTGTGATCAGAGAGTGCCGCTCCAAACAAAATACGCGC 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTTCCAAGCCTCTAAAAATGTGCTTTGTGATCAGAGAGTGCCGCTCCAAACAAAATACGCGC 720

Qy     721 GCTGCCTTTCGAGGCCAGTGAGCTCAGCCTCCAAGGCTTTAAAGCCACATTTCAGCAAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GCTGCCTTTCGAGGCCAGTGAGCTCAGCCTCCAAGGCTTTAAAGCCACATTTCAGCAAG 780

Qy     781 AGAAAGCGCTGAGAGCTCGCAGGTTCATTGAAGAAGGGCAAAGCACTGGTTTCTCTCCTTA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AGAAAGCGCTGAGAGCTCGCAGGTTCATTGAAGAAGGGCAAAGCACTGGTTTCTCTCCTTA 840

Qy     841 GAAAAGTAGGTTCTTGGCTTGAGTAGACTGGGCTTGCTTGCTTTAGTTTTTAGTTGAAGGGAAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GAAAAGTAGGTTCTTGGCTTGAGTAGACTGGGCTTGCTTGCTTTAGTTTTTAGTTGAAGGGAAT 900

Qy     901 GTACGTAAAAACAAAAATAGGGGCTTGGGCTGGTCAAAAGGAAGAGAACAAGCAGGAGTGGGATGGATGGA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GTACGTAAAAACAAAAATAGGGGCTTGGGCTGGTCAAAAGGAAGAGAACAAGCAGGAGTGGGATGGATGGA 960
```

```
Qy     961 TGGATGGATGGATGTGTATGGATAGATAGATAGATGATGGTGGTGTTGGCATGGAAGGTCGCAGC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TGGATGGATGGATGTGTATGGATAGATAGATAGATGATGGTGGTGTTGGCATGGAAGGTCGCAGC 1020

Qy    1021 ACAAAACCAAAGCTGATTGGAAACAATTAATTGTGGGGTGCTCGAGGGGGGAAGGTCGGAGCC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ACAAAACCAAAGCTGATTGGAAACAATTAATTGTGGGGTGCTCGAGGGGGGAAGGTCGGAGCC 1080

Qy    1081 TTTGGGGCAGCTTTGAGAAGCGGTACAAGAGTTCGTGCCTGTGTCCAGCCTGGAGCC 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TTTGGGGCAGCTTTGAGAAGCGGTACAAGAGTTCGTGCCTGTGTCCAGCCTGGAGCC 1140

Qy    1141 AGCCAGTGCATTTATTTAAGCTCTTAGAAGCAACTCCTTGCCCCAGGAATGCGTGACCC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 AGCCAGTGCATTTATTTAAGCTCTTAGAAGCAACTCCTTGCCCCAGGAATGCGTGACCC 1200

Qy    1201 CTGAGAATGGGGTCCACGCAATCTCTCACACTTCCTTCTTCTCCGTGGGGGGATACTGGACTCGTG 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTGAGAATGGGGTCCACGCAATCTCTCACACTTCCTTCTTCTCCGTGGGGGGATACTGGACTCGTG 1260

Qy    1261 CCTCTCGCGCCCCATTCTTCTTCACGCATATCCATGACTTAAATTTCACTTTCTGATCAC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CCTCTCGCGCCCCATTCTTCTTCACGCATATCCATGACTTAAATTTCACTTTCTGATCAC 1320

Qy    1321 GGTACGTCCATAAAGCCAGTATTACACTTAAATGAAGGTATTCTTTTTGTAATCGTTTTT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GGTACGTCCATAAAGCCAGTATTACACTTAAATGAAGGTATTCTTTTTGTAATCGTTTTT 1380

Qy    1381 TTTAGAAGGTAAACAAATTTAATAAAGCTACCAATATATGAGAAAAAAAAAAAAAA 1439
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TTTAGAAGGTAAACAAATTTAATAAAGCTACCAATATATGAGAAAAAAAAAAAAAA 1439
```

```
RESULT 2
BC032745
LOCUS        BC032745               2310 bp    mRNA    linear    HTC  04-MAR-2003
DEFINITION   Homo sapiens, DKFZp434N161 protein, clone IMAGE:5574690, mRNA.
ACCESSION    BC032745
VERSION      BC032745.1  GI:22749759
KEYWORDS     HTC.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 2310)
  AUTHORS    Strausberg,R.
  TITLE      Direct Submission
  JOURNAL    Submitted (06-JUN-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
  REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: National Institutes of Health Intramural
             Sequencing Center (NISC),
             Gaithersburg, Maryland;
             Web site: http://www.nisc.nih.gov/
             Contact: nisc_mgc@nhgri.nih.gov
             Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
             Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
             Dietrich,N.L., Granite,S.J., Guan,X., Gupta,J., Haghighi,P.,
             Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
             Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
             McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
             Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
             Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2003, 05:31:01 / Search time 321.205 Seconds
(without alignments)
10025.973 Million cell updates/sec

Title: US-09-818-143-6
Perfect score: 1439
Sequence: 1 gagtatcccttgtttaatca..........agaaaaaaaaaaaaaaaaaa 1439

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

1 number of hits satisfying chosen parameters: 3008958

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Database :
Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	Match ID	Description
1	1439	100.0	1439	9	US-09-818-143-6	Sequence 6, Appli
2	1439	100.0	1439	10	US-09-974-298-148	Sequence 148, App
3	1257.6	87.4	1324	14	US-10-125-237-4	Sequence 4, Appli
4	1257.6	87.4	1324	14	US-10-105-991-4	Sequence 4, Appli
5	629.8	43.8	633	10	US-09-864-864-159	Sequence 159, App
6	562	39.1	562	14	US-09-918-995-427	Sequence 427, App
7	405.8	28.2	495	11	US-09-918-995-22094	Sequence 22094, A
8	386.4	26.9	565	11	US-09-833-381-1144	Sequence 1144, Ap
9	386	26.8	386	11	US-09-822-846-30	Sequence 30, Appl
10	355.4	24.7	396	10	US-09-918-995-7821	Sequence 7821, Ap
11	300.4	20.9	309	10	US-09-954-456-1950	Sequence 1950, Ap
12	242.2	16.8	268	10	US-09-833-381-1143	Sequence 1143, Ap
13	188.6	13.1	303	10	US-09-783-590-1803	Sequence 1803, Ap
14	180.2	12.5	297	10	US-09-783-381-1935	Sequence 1935, Ap
15	129.8	9.0	517	10	US-09-833-381-1145	Sequence 1145, Ap
16	118	8.2	196	10	US-09-783-590-4264	Sequence 4264, Ap

c	17	50.4	3.5	487	13	US-10-027-632-77534	Sequence 77534, A
c	18	49.4	3.4	40267	11	US-09-978-244A-25	Sequence 25, Appl
c	19	48.2	3.3	1503841	9	US-09-795-688-1	Sequence 1, Appli
c	20	48.2	3.3	1503841	9	US-09-795-608-1	Sequence 1, Appli
c	21	48.2	3.3	1503841	10	US-09-946-007-1	Sequence 1, Appli
c	22	48	3.3	32189	9	US-09-764-878-379	Sequence 379, App
c	23	48	3.3	32189	14	US-10-079-854-379	Sequence 379, App
c	24	48	3.3	32189	9	US-09-764-878-377	Sequence 377, App
c	25	48	3.3	32221	14	US-10-079-854-377	Sequence 377, A
c	26	48	3.3	32221	13	US-10-027-632-14744	Sequence 14744, A
c	27	47.8	3.3	518	13	US-10-027-632-14745	Sequence 14745, A
c	28	47.8	3.3	518	13	US-10-027-632-14746	Sequence 14746, A
c	29	47.8	3.3	518	13	US-10-027-632-14747	Sequence 14747, A
c	30	47.6	3.3	7597	12	US-10-311-455-985	Sequence 985, App
c	31	47.6	3.3	32646	12	US-10-017-161-1657	Sequence 1657, Ap
c	32	47.2	3.3	483	13	US-10-027-632-40348	Sequence 40348, A
c	33	47	3.3	11832	9	US-09-037-657-38	Sequence 38, Appli
c	34	47	3.3	30175	9	US-09-738-878-3	Sequence 3, Appli
c	35	47	3.3	30175	13	US-10-163-381-3	Sequence 3, Appli
c	36	46.8	3.3	1006	13	US-10-027-632-273001	Sequence 273001,
c	37	46.8	3.3	68	11	US-09-969-708-464	Sequence 464, App
c	38	46.2	3.2	591	13	US-10-027-632-130534	Sequence 130534,
c	39	46.2	3.2	725	13	US-10-027-632-243438	Sequence 243438,
c	40	46	3.2	3768	10	US-09-952-903C-17	Sequence 17, Appl
c	41	46	3.2	3768	10	US-09-954-531-164	Sequence 164, App
c	42	46	3.2	3768	10	US-09-954-531-373	Sequence 373, App
c	43	46	3.2	3768	12	US-09-873-319-689	Sequence 689, App
c	44	46	3.2	3768	12	US-09-960-706-1041	Sequence 1041, Ap
c	45	46	3.2	32816	9	US-09-729-094-3	Sequence 3, Appli

ALIGNMENTS

Query Match   100.0%; Score 1439; DB 9; Length 1439;
Best Local Similarity   100.0%;
Matches 1439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GAGTATCCCTTGTTTAATCACTTTTTGTGGTTAAAAGAGAGACCTTTGGGTCAGTCTGCCTCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GAGTATCCCTTGTTTAATCACTTTTTGTGGTTAAAAGAGAGACCTTTGGGTCAGTCTGCCTCA 60

Qy   61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTTTCACTCTATTTCTTCTCGTCTCAAGAA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTTTCACTCTATTTCTTCTCGTCTCAAGAA 120

Qy   121 AGAAGAAAAAAGAGACAAAATTACCCAGAAAACCCCTCCCCACATGGAGGCCTTGG 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGAAGAAAAAAGAGACAAAATTACCCAGAAAACCCCTCCCCACATGGAGGCCTTGG 180

Qy   181 CAAATGTTAATTTCCTTAGAAAAATCCTTCAGACCTGAAGACGCCAGGAAAAGAATCTGGCT 240

RESULT 2
US-09-974-298-148
; Sequence 148, Application US/09974298
; Patent No. US2002015623A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 148
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2002015623A1 1556751CB1
US-09-974-298-148

Query Match   100.0%; Score 1439; DB 10; Length 1439;
Best Local Similarity   100.0%;
Matches 1439; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GAGTATCCCTTGTTTAATCACTTTTGTGGTTAAAAGAGAGACCTTTGGGTCAGTCAGTTCGGCCTCA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GAGTATCCCTTGTTTAATCACTTTTGTGGTTAAAAGAGAGACCTTTGGGTCAGTCAGTTCGGCCTCA 60

Qy   61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTCACTCTATTCCTTCCCTGTCTCAAGAA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTCACTCTATTCCTTCCCTGTCTCAAGAA 120

Qy 121 AGAGAAAAAGAGACAAATTACCCAGAAAACCCTCCCTTCCCCACATGGAGGCCTTGG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGAGAAAAAGAGACAAATTACCCAGAAAACCCTCCCTTCCCCACATGGAGGCCTTGG 180

Qy 181 CAAATGTTAATTTTCCTAGAAAATCCTTCAGACGTTGAAGACGCAGGAAATCTGGCT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CAAATGTTAATTTTCCTAGAAAATCCTTCAGACGTTGAAGACGCAGGAAATCTGGCT 240

Qy 241 CTCAGGGTGGCTTCTCGCGTTCCCGCGCGGCCCAGGCCCCAGACTATGGTCACAGGGGCCGTCCT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CTCAGGGTGGCTTCTCGCGTTCCCGCGCGGCCCAGGCCCCAGACTATGGTCACAGGGGCCGTCCT 300

Qy 301 GTTCCTCCCCGGGACTCCAGAATTTCTCTCCGCGCCAAAGGAAAAGAAAACAGGGCATGCGCTT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 GTTCCTCCCCGGGACTCCAGAATTTCTCTCCGCGCCAAAGGAAAAGAAAACAGGGCATGCGCTT 360

Qy 361 GTTGGCAAAACGCCAGGCGGCGGCCCGGCTCCCAAAAACCCATGTGTGTACGATTAAAAAGTTGGCC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 GTTGGCAAAACGCCAGGCGGCGGCCCGGCTCCCAAAAACCCATGTGTGTACGATTAAAAAGTTGGCC 420

Qy 421 GTCCCCAAGGCCTTCCCAGCCGCAAACTTAAAGACGACAGGGCTTTGCTGAAAACAATGG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GTCCCCAAGGCCTTCCCAGCCGCAAACTTAAAGACGACAGGGCTTTGCTGAAAACAATGG 480

Qy 481 GCCAGCTGGGCTTTTAACAACCTTAGAAGACTTTCCCGGAGCTGCCTGGAACAGAGCCTGCG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 GCCAGCTGGGCTTTTAACAACCTTAGAAGACTTTCCCGGAGCTGCCTGGAACAGAGCCTGCG 540

Qy 541 GGAAAACGGGGCTTGCCAGAGACACTCACAGTTGCCAGAGACACTAGTTCCTCCCTAA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GGAAAACGGGGCTTGCCAGAGACACTCACAGTTGCCAGAGACACTAGTTCCTCCCTAA 600

Qy 601 GAATCTCCACATCATTGTCTTTGTGCCTTTTTTCCTTGGCTTTCCCGGAACACAGAAAGGGAAGG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 GAATCTCCACATCATTGTCTTTGTGCCTTTTTTCCTTGGCTTTCCCGGAACACAGAAAGGGAAGG 660

Qy 661 GTTCCAAGCCTCTAAAAAATGTGCTTTTGTGATCAGAGGTGGCTCCAAACCAAATACGGCGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GTTCCAAGCCTCTAAAAAATGTGCTTTTGTGATCAGAGGTGGCTCCAAACCAAATACGGCGC 720

Qy 721 GCTGCCCTTTCGAGGCCAGTGAGCTCAGCCTTCCAAGGCTTTAAAGCCCACATTTCAGCAAG 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 GCTGCCCTTTCGAGGCCAGTGAGCTCAGCCTTCCAAGGCTTTAAAGCCCACATTTCAGCAAG 780

Qy 781 AGAAAGCGCTGAAGACGCTCGGCAGGTTCATTAAAGRAAGGCAAAGCACTGGTTTCTCTCCTTA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 AGAAAGCGCTGAAGACGCTCGGCAGGTTCATTAAAGAAAGGCAAAGCACTGGTTTCTCTCCTTA 840

Qy 841 GAAAAGTAGGTTTCTTGGCTTGAATGTAGACTGGGCTTGTGATTTTTAGTGAAGGGAAT 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 GAAAAGTAGGTTTCTTGGCTTGAATGTAGACTGGGCTTGTGATTTTTAGTGAAGGGAAT 900

Qy 901 GTAGGTAAAACAAAATAGGGCTTGGCTGGTCGTCAAAGAAGACCAGGATGGATGGA 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 GTAGGTAAAACAAAATAGGGCTTGGCTGGTCGTCAAAGAAGACCAGGATGGATGGA 960

Qy 961 TGGATGGATGGGATGTAGGATAGAATGGTGTTTTGCATGTAAATTGCAGAGAA 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 TGGATGGATGGGATGTAGGATAGAATGGTGTTTTGCATGTAAATTGCAGAGAA 1020

Qy 1021 ACAAAAACCAAAGCTGATTGGTGGAAACAATTAATTGTGGGTGTGAAGGTTCGCAGC 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACAAAAACCAAAGCTGATTGGTGGAAACAATTAATTGTGGGTGTGAAGGTTCGCAGC 1080

Qy 1081 TTTGGGCAGCTTTGAGAAGCGGGTACAAGAGTTCTGTGCCTGTGTGTTCAGCCTGGAGCC 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TTTGGGCAGCTTTGAGAAGCGGGTACAAGAGTTCTGTGCCTGTGTGTTCAGCCTGGAGCC 1140

Qy 1141 AGCCAGTGCATTTATTTATTTAAGCTTAGAGAGCAACTCCTTGCCCTTGGCCTGACC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 AGCCAGTGCATTTATTTATTTAAGCTTAGAGAGCAACTCCTTGCCCTTGGCCTGACC 1200

Qy 1201 CTGAGATGGGTCCACCGCATCTCTCACACTTCCTCTCTCCGTGGATACTGGACTCGTG 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGAGATGGGTCCACCGCATCTCTCACACTTCCTCTCTCCGTGGATACTGGACTCGTG 1260

Qy 1261 CCTCTGGGCCCATTCTCTCTCACGCAATATCCATGAGCTTAATTCACTTTCTGATCAC 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 CCTCTGGGCCCATTCTCTCTCACGCAATATCCATGAGCTTAATTCACTTTCTGATCAC 1320

Qy 1321 GGTACGTCCATAAAGCCAGTATTACACTTAAATGAAGTAATCGTTTTT 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 GGTACGTCCATAAAGCCAGTATTACACTTAAATGAAGTAATCGTTTTT 1380

Qy 1381 TTTAGAAGGTAAACAAATTTAATAAAGTACCAAATAATGAGAAAAAAAAAAAAA 1439
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 TTTAGAAGGTAAACAAATTTAATAAAGTACCAAATAATGAGAAAAAAAAAAAAA 1439

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 13, 2003, 03:56:05 ; Search time 86.2666 Seconds
(without alignments)
7362.646 Million cell updates/sec

Title: US-09-818-143-6
Perfect score: 1439
Sequence: 1 gagtatccctgtttaatca.........agaaaaaaaaaaaaaaaaaa 1439

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

1 number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptcdata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptcdata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptcdata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptcdata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptcdata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptcdata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1264	87.8	4772	4	US-09-484-970B-133	Sequence 133, App
2	60.6	4.2	7218	4	US-08-232-463-14	Sequence 14, Appl
3	45.4	3.2	374	2	US-08-329-799-34	Sequence 34, Appl
4 c	44.4	3.2	152331	3	US-09-128-155-16	Sequence 130, App
5	44.4	3.1	176373	3	US-09-128-155-17	Sequence 130, App
6	44.2	3.1	1315	3	US-09-721-82zA-10	Sequence 130, App
7	43.8	3.0	367	3	US-09-332-766A-9	Sequence 10, Appl
8	43.8	3.0	857	2	US-08-460-040-1	Sequence 9, Appli
9	43.6	3.0	1172	1	US-07-945-288-9	Sequence 9, Appli
10	43.6	3.0	1172	1	US-08-462-831-9	Sequence 9, Appli
11	43.6	3.0	1172	1	US-08-461-809-9	Sequence 9, Appli
12	43.6	3.0	1172	1	US-08-461-441-9	Sequence 9, Appli
13	43.6	3.0	1172	5	PCT-US93-08518-9	Sequence 9, Appli
14	43.4	3.0	7218	1	US-08-232-463-14	Sequence 14, Appl
15	43	3.0	3693	1	US-09-056-226-1	Sequence 1, Appli
16	42.6	3.0	3207	1	US-09-056-226-1	Sequence 1, Appli
17	42.6	3.0	3207	1	US-08-483-332-1	Sequence 1, Appli
18	42.6	3.0	3207	2	US-08-478-882-1	Sequence 1, Appli
19	42	2.9	3207	1	PCT-US93-06251-23	Sequence 23, Appl
20	41.8	2.8	11558	5	US-09-741-154-3	Sequence 3, Appli
21	40.6	2.8	16389	4	US-09-369-247-19	Sequence 19, Appl
22	40	2.8	2089	4	US-09-369-247-19	Sequence 19, Appl
23	39.6	2.8	32654	4	US-09-385-982-155	Sequence 155, App
24	39.6	2.8	450	4	US-09-801-191A-3	Sequence 3, Appli
25	39.6	2.8	3347	4	US-09-702-705-318	Sequence 318, App
26 c	39.6	2.7	957	4	US-09-736-457-318	Sequence 318, App
27	39.2	2.7	1091	4	US-09-328-965-1	Sequence 1, Appl
28	39.2	2.7	1474	3	US-08-821-994-64	Sequence 64, Appl
29	39.2	2.7	1493	6	5340934-5	Patent No. 5340934
30 c	39.2	2.7	3850	3	US-08-329-799-34	Sequence 34, Appl
31	38.8	2.7	855	3	US-08-906-765-130	Sequence 130, App
32	38.8	2.7	855	3	US-08-906-616-130	Sequence 130, App
33	38.8	2.7	855	3	US-08-639-075A-130	Sequence 130, App
34	38.8	2.7	855	3	US-09-012-431-130	Sequence 130, App
35	38.8	2.7	855	4	US-09-012-692-130	Sequence 130, App
36	38.8	2.7	855	4	US-08-906-613-130	Sequence 130, App
37	38.8	2.7	1190	4	US-09-390-207-1	Sequence 1, Appli
38	38.8	2.7	19124	2	US-08-487-8268-13	Sequence 13, Appl
39	38.6	2.7	7859	1	US-07-854-596B-4	Sequence 4, Appli
40	38.6	2.7	7859	3	US-08-450-905B-15	Sequence 15, Appl
41	38.6	2.7	7859	3	US-07-982-759P-15	Sequence 15, Appl
42	38.4	2.7	1378	3	US-09-149-476-208	Sequence 208, App
43	38.4	2.7	2323	4	US-09-149-476-24	Sequence 24, Appl
44 c	38.4	2.7	4072	3	US-09-272-496-7	Sequence 7, Appli
45 c	38.2	2.7	711	3	US-08-998-416-786	Sequence 786, App

ALIGNMENTS

RESULT 1
US-09-484-970B-133

; Sequence 133, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
;
; SEQ ID NO 133
; LENGTH: 4772
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 245487.3CB1
; NAME/KEY: unsure
; LOCATION: 3196-3488
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-133

Query Match 87.8%; Score 1264; DB 4; Length 4772;
Best Local Similarity 99.5%; Pred. No. 0; Mismatches 5; Indels 1; Gaps 1;
Matches 1278; Conservative 0;

```
Qy   144 CCAGAAACCCCTCCCTTCCCACATGGAGGCCTTGGCAAATGTTAATTTTTCCTTAGAAAA 203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3489 CCAGAAACCCCTCCCTTCCCACATGGAGGCCTTGGCAAATGTTAATTTTTCCTTAGAAAA 3548

Qy   204 TCCTTCAGACTTGAAGAGCACAGAAAAGAAATCTGGCTCTCCAGGGTGGCTTGCGTCCCC 263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3549 TCCTTCAGACCTGAAGAGCGCAGAAAAGAAATCTGGCTCTCCAGGGTGGCTTGCGTCCCC 3608

Qy   264 GCGCGCCAGGCCCCAGACTTATGGTTCACAGGGCCCGTCCTGTTCCTCCCCCGGGAAT 323
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3609 GCGGCCAGGCCCCAGACTATGGTTCACAGGGCCCGTCCTGTTCCTCCCCCGGGACTCCAGAAT 3668

Qy   324 TTCTCTCTCCTCAAAGGAAGAAAAACACGGGCCATGGCCGTTCGTTGTTGGCCAAAACGCAA 383
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3669 TTCTCTCTCCTCAAAGGAAGAAAAACAGGGCCATGGCGCTTGTTGTTGGCCAAAACGCAGGGGCT 3728

Qy   384 CCCAAAAACCCCATGTGTGTACGGATTAAAAGTTTGGCCGTCCCCAGGGCCTCCCAGGGCAAA 443
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3729 CCCAAAAACCCCATGTGTGTACGGATTAAAAGTTTGGCCGTCCCCAGGGCCTCCCAGGGCAAA 3788
```

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367

GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: DORNER, P.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pT2gpt-F1s
US-08-232-463-14

Query Match   4.2%; Score 60.6; DB 1; Length 7218;
Best Local Similarity   7.5%; Pred. No. 5.3e-07;
Matches 30; Conservative 211; Mismatches 160; Indels 0; Gaps 0;

Qy   775 AGCAAGAGAAAGCGCTGAGAGCTCGCAGGTTCATTAAGAAGAGGAAAGCACTGGTTCTC 834
          ::::::::::::: :::::: :::::
Db   1406 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1347

Qy   835 TCCTTAGAAAGTAGGTTCTTCGGCTTGATGTAGAGACTGGCTGTTGCTTTGATTTTAGTGAA 894
          :::::::::::::::::::::::::::::
Db   1346 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1287

Qy   895 GGGAATGTACGTAAAACAAAATAGGGCTTGGCGTCAAAGGAGACAAGCAGGATGGATG 954
          :::::::::::::::::::::::::::
Db   1286 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1227

Qy   955 GATGGATGGAATGGAATGGATGTATGGACAGAATAGATAGATAGGTGTTTGCATGTAAATTGCA 1014
          :::::::::::::::::::::::::::::
Db   1226 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1167

Qy   1015 GAGAAAACAAAACCAAAGCTGATTGGAAACAATTAATTGTGGGTGTCTCGAGGGGAAGGT 1074
          :::::::::::::::::::::::::::::
Db   1166 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1107

Qy   1075 CGCAGCTTTGGGCAGCTTTGAGAACCGGTACAAGAGTTCTCGTGCCTGTGTGTCCAGCCT 1134
          ::::::::::::::::::::::::::
Db   1106 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGGAAGCTCCCTCGACCT 1047

OM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: August 13, 2003, 02:51:00 ; Search time 380.307 seconds
                                          (without alignments)
                                          10214.090 Million cell updates/sec

Title: US-09-818-143-6
Perfect score: 1439
Sequence: 1 gagtacccctgttaatca.........agaaaaaaaaaaaaaaaa 1439

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
    1 number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_19Jun03:*
          1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
          2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
          3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
          4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
          5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
          6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
          7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
          8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
          9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
         10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
         11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
         12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
         13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
         14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
         15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
         16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
         17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
         18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
         19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
         20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
         21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
         22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
         23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
         24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
         25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1439	100.0	1439	21	AAZ52354	NSEQ gene-6 associ
2	1439	100.0	1439	24	AAL42450	Human matrix-remod
3	1439	100.0	1439	25	ABX77619	Differentially exp
4	1326	92.1	3734	22	AAS33201	DNA encoding human
5	1317.8	91.6	5487	25	ACC46580	Human dithp cytosk
6	1264	87.8	4772	25	ABS70476	Human bone remodel
7	1263.6	87.8	5859	24	ABS67329	Breast specific pol
8	629.8	43.8	633	24	ABX09622	Human ovarian tumo

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

RESULT 1
AAZ52354

ID   AAZ52354 standard; cDNA; 1439 BP.
XX
AC   AAZ52354;
XX
DT   24-JUL-2000 (first entry)
XX
DE   NSEQ gene-6 associated with matrix remodelling.
XX
KW   NSEQ gene; matrix-remodeling gene; Incyte clone 1556751; cancer;
KW   matrix-remodeling disease; cardiomyopathy; arthritis; angiogenesis;
KW   diabetic necrosis; atherosclerosis; fibrosis; ulceration; cytostatic;
KW   cardioactive; antiarthritic; angiogenic; antiarteriosclerotic;
KW   anticancer; ss.
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   168..467
FT   /*tag= a
FT   /product= "PSEQ protein"
XX
PN   WO20002198G-A2.
XX
PD   20-APR-2000.
XX
PF   06-OCT-1999; 99WO-US023315.
XX
PR   09-OCT-1998; 98US-0169289.

ALIGNMENTS

9	562	39.1	562	22	AAH33361	Human colon cancer
10	386	26.8	386	24	ABK35639	cDNA sequence #30
11	370.4	25.7	397	20	AAV86538	EST clone AU118
12	300.4	20.9	309	24	ABL62265	Colon adenocarcino
13	300.4	20.9	309	24	ABL66640	Colon adenocarcino
14	115.4	8.0	89328	16	ABL61995	Lung cancer relate
15	104.2	7.2	111	16	AAT22354	Human gene signatu
16	94.6	6.6	5487	25	ACC46580	Human dithp cytosk
17	51	3.5	51	22	AAH90355	Human clone cg4394
18	50.6	3.5	5675571	25	AAK70122	Human immune/haema
19	50.4	3.5	40267	24	AAD53224	Human chromosome 3
20	49.4	3.4	23885	22	AAS18958	Human cadherin-11k
21	48.2	3.3	23885	22	AAK70103	Human immune/haema
22	48.2	3.3	23885	22	AAK73173	Human immune/haema
23	48.2	3.3	1503841	21	ABT00010	Human neuregulin 1
24	48.2	3.3	1503841	24	ABT01503	Human neuregulin 1
25	48.2	3.3	1503900	22	AAK95240	Human neuregulin-1
26	48	3.3	1503900	22	AAK96733	Human neuregulin-1
27	48	3.3	32189	21	AAC17114	Human secreted pro
28	48	3.3	32189	22	AAS30115	Human lung antigen
29	48	3.3	32221	22	AAS30114	Human lung antigen
30	47.6	3.3	7597	24	AAS30113	Human immune syste
31	47.2	3.3	523	22	ABL33012	Human immune syste
32	47	3.3	11832	22	ABA13989	Human nervous syst
33	47	3.3	11832	19	AAV27148	Nucleotide sequenc
34	47	3.3	11832	22	AAD04198	Murine NR6 genomic
35	47	3.3	30175	24	AAD32742	Human drug metabol
36	47	3.3	240825	24	AAP24497	Human PG-3 gene.
37	46.2	3.2	240825	22	ABQ81802	Human PG-3 gene SE
38	46.2	3.2	68	24	ABK24292	MM211 microsatelli
39	46.2	3.2	438	24	ABS73251	Human translocatio
40	46.2	3.2	5460	24	ABN80044	Human chemically m
41	46.2	3.2	117574	24	AAL45288	Human KCNB1 gene.
42	46	3.2	213456	22	AAL54072	Genomic DNA encodi
43	46	3.2	3768	24	ABK64794	Human benign post
44	46	3.2	3768	24	ABL62894	Breast cancer rela
45	46	3.2	3768	24	ABL63103	Breast cancer rela
46	46	3.2	3768	24	ABL69598	Prostate cancer re

```
PA  (INCY-) INCYTE PHARM INC.
XX
PI  Walker MG, Volkmuth W, Klingler TM;
XX
DR  WPI; 2000-317934/27.
DR  P-PSDB; AAY70744.
XX
PT  Protein co-expressed with matrix-remodeling proteins, useful in the
PT  diagnosis and treatment of cancer, cardiomyopathy, arthritis,
PT  angiogenesis, diabetic necrosis, atherosclerosis, fibrosis, and
PT  ulceration .
XX
PN  Claim 1; Page 40-41; 55pp; English.
XX
PA  The present sequence is NSEQ gene that is co-expressed with one or more
CC  known matrix-remodeling genes in a number of biological samples using an
CC  expression vector. This sequence was identified from the
CC  Incyte clone 1556751. The gene, protein, and antibody sequences can be
CC  used in the diagnosis, and treatment or prevention of a disease
CC  associated with its altered expression. The diseases that can be treated
CC  are matrix-remodeling diseases, including cancer, cardiomyopathy,
CC  arthritis, angiogenesis, diabetic necrosis, atherosclerosis, fibrosis,
CC  and ulceration.
XX
SQ  Sequence 1439 BP; 392 A; 336 C; 338 G; 373 T; 0 other;

Query Match           100.0%;  Score 1439;  DB 21;  Length 1439;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1439;  Conservative   0;  Mismatches    0;  Indels   0;  Gaps   0;

Qy        1 GAGTATCCCTTGTTTAATCACTTTTGTGGTTAAAAGAGAGAACCTTGGGTTCAGTCTCCCTCA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GAGTATCCCTTGTTTAATCACTTTTGTGGTTAAAAGAGAGAACCTTGGGTTCAGTCTCCCTCA 60

Qy       61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTTCACTCTATTTCTTCCCTGTCTCCAAGAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTTCACTCTATTTCTTCCCTGTCTCCAAGAA 120

Qy      121 AGAAGAAAAAAGAGAGACAAAATTACCCAGAGAAACCCTCTCCTTCCCACATGGAGGGCTTGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AGAAGAAAAAAGAGAGACAAAATTACCCAGAGAAACCCTCTCCTTCCCACATGGAGGGCTTGG 180

Qy      181 CAAATGTTAATTTTCCTAGAAAAATCCTTCAGACGCAGGGAGAAAGAATCTGGCT 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CAAATGTTAATTTTCCTAGAAAAATCCTTCAGACGCAGGGAGAAAGAATCTGGCT 240

Qy      241 CTCAGGGTGGCTTCTCGGTCCCCGCGGCCCCAGACCTATGGGTCACAGGGCGGTCCT 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CTCAGGGTGGCTTCTCGGTCCCCGCGGCCCCAGACCTATGGGTCACAGGGCGGTCCT 300

Qy      301 GTTCCTTCCCCGGGACTCCAGAATTTCTCTCCTCAAAAGGAAAAAAAACCCATGCCGCTT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GTTCCTTCCCCGGGACTCCAGAATTTCTCTCCTCAAAAGGAAAAAAAACCCATGCCGCTT 360

Qy      361 GTTGGCAAAACGACAGGGCGGCGGCGGTCCCAAAAACCCCATGTGGTACGAATTAAAAGTTGGCC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GTTGGCAAAACGACAGGGCGGCGGCGGTCCCAAAAACCCCATGTGGTACGAATTAAAAGTTGGCC 420

Qy      421 GTCCCCAGCGCCTCCCACGCGCAAACTTAAAGAGACAGGGCCTTTGCTGAAAACAAACATGG 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GTCCCCAGCGCCTCCCACGCGCAAACTTAAAGAGACAGGGCCTTTGCTGAAAACAAACATGG 480

Qy      481 GCCAGCTGGGCTTTTTAACAACCTAGACTTTCCGGAGCTGCCTGGACAGGGCCCTTAA 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 GCCAGCTGGGCTTTTTAACAACCTAGACTTTCCGGAGCTGCCTGGACAGGGCCCTTAA 540

Qy      541 GGAAACGGGGCTTGCAGAGACACTCACAGTTCCTTCGTTTGGTTCCCCTAA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GGAAACGGGGCTTGCAGAGACACTCACAGTTCCTTCGTTTGGTTCCCCTAA 600

Qy      601 GAATCTCCACATCATTGTCTTTCTTGTGCCTTTTCCTTGGTGAGCAACAGAAGAAGGGGAAGG 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GAATCTCCACATCATTGTCTTTCTTGTGCCTTTTCCTTGGTGAGCAACAGAAGAAGGGGAAGG 660

Qy      661 GTTCCAAGCCTTCTAAAAAATGTTTGTGATCAGGAGTCGGCTCCAAACCAAAATACGGCC 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GTTCCAAGCCTTCTAAAAAATGTTTGTGATCAGGAGTCGGCTCCAAACCAAAATACGGCC 720

Qy      721 GCTGCCCCTTTCGAGGCGGCCAGTGAGCTCAGCCTCCAAGGCGCTTTAAAGCCACATTTCAGCAAG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GCTGCCCCTTTCGAGGCGGCCAGTGAGCTCAGCCTCCAAGGCGCTTTAAAGCCACATTTCAGCAAG 780

Qy      781 AGAAAGGCGCTGAGAGCTTCGCAGGTTCATTAAAGAAGGCAAAACCACTGGGTTTCTCCCTTA 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 AGAAAGGCGCTGAGAGCTTCGCAGGTTCATTAAAGAAGGCAAAACCACTGGGTTTCTCCCTTA 840

Qy      841 GAAAAGTAGGTTCTTCGGCTTGATGTAGACTGGCTTGCTTTGGATTTTTAGTGAAGGGAAT 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GAAAAGTAGGTTCTTCGGCTTGATGTAGACTGGCTTGCTTTGGATTTTTAGTGAAGGGAAT 900

Qy      901 GTACGTAAAACAAAATAGGGCTTGGGCTGGTTCAAAGGAGCAAGCAGGAGGGGATGGATGGA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GTACGTAAAACAAAATAGGGCTTGGGCTGGTTCAAAGGAGCAAGCAGGAGGGGATGGATGGA 960

Qy      961 TGGGATGGAGTGATGGAGTATGGAGTGAATAGAGATAGATAGAAGGTGGTTGCATTGTAAATTGCAGAA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TGGGATGGAGTGATGGAGTATGGAGTGAATAGAGATAGATAGAAGGTGGTTGCATTGTAAATTGCAGAA 1020

Qy     1021 ACAAAAACCAAAAGCTGATTGGAAACAAATTAATTGTGGTGTCTGAGGGGGAGGTCGGAGC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 ACAAAAACCAAAAGCTGATTGGAAACAAATTAATTGTGGTGTCTGAGGGGGAGGTCGGAGC 1080

Qy     1081 TTTGGGCACGTTTGAGAAGCGGTACAAGAGTTCGTGCGCCGTGTGTCCAGGCGGTCCGGAGCC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TTTGGGCACGTTTGAGAAGCGGTACAAGAGTTCGTGCGCCGTGTGTCCAGGCGGTCCGGAGCC 1140

Qy     1141 AGCCAGTGCATTTATTTTAAGCTTAGAAGCAACTCCCTTCGGCCCAGGAATGGGTGACCC 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 AGCCAGTGCATTTATTTTAAGCTTAGAAGCAACTCCCTTCGGCCCAGGAATGGGTGACCC 1200

Qy     1201 CTGAGATGGGTCCACGCATCTCTTCTACACTTCCTTCTCTCTCCGTGGGGGATACGGACTCGTG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 CTGAGATGGGTCCACGCATCTCTTCTACACTTCCTTCTCTCTCCGTGGGGGATACGGACTCGTG 1260

Qy     1261 CCTCTGCCCCATTCTCTTCTCCGCAATCCATGAGCTTTAAATTCACTTTCTGATCAC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CCTCTGCCCCATTCTCTTCTCCGCAATCCATGAGCTTTAAATTCACTTTCTGATCAC 1320

Qy     1321 GGTACGTCCATAAAAGCCAGTAATTACACTTAAATGAAGTATTCTTTTTTGTAATCGTTTTT 1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 GGTACGTCCATAAAAGCCAGTAATTACACTTAAATGAAGTATTCTTTTTTGTAATCGTTTTT 1380

Qy     1381 TTTAGAAGGTAAACAAAATTTAATAAAGCTACCAATAATGAGAAAAAAAAAAAAAAAA 1439
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 TTTAGAAGGTAAACAAAATTTAATAAAGCTACCAATAATGAGAAAAAAAAAAAAAAAA 1439
```

```
RESULT 2
AAL42450
ID   AAL42450 standard; DNA; 1439 BP.
XX
AC   AAL42450;
XX
DT   11-JUL-2002  (first entry)
XX
DE   Human matrix-remodeling-associated nucleotide 6.
XX
KW   Human; d6; matrix-remodeling gene; extracellular matrix; gene;
KW   matrix-remodeling-associated nucleotide; screening;
KW   matrix remodelling-associated disease; angiogenesis; arthritis;
KW   atherosclerosis; cancer; cardiomyopathy; diabetic necrosis; fibrosis;
KW   ulceration.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
```

```
FT   CDS              168..467
FT                    /*tag= a
FT                    /product= "Human matrix-remodeling-associated protein 2"
XX
PN   US2002019000-A1.
XX
PD   14-FEB-2002.
XX
PF   26-MAR-2001; 2001US-0818143.
XX
PR   09-OCT-1998; 98US-0169289.
XX
PI   Walker MG, Volkmuth W, Klingler TM;
PA   (WALK/) WALKER M G;
PA   (VOLK/) VOLKMUTH W.;
PA   (KLIN/) KLINGLER T M.
XX
PT   New isolated polynucleotide coexpressed with matrix-remodeling genes,
PT   useful in diagnosis, prognosis, prevention and treatment of diseases
PT   associated with matrix-remodeling such as angiogenesis, arthritis and
PT   cancer.
XX
PS   Claim 1; Page 21; 63pp; English.
XX
CC   The invention comprises human nucleotide sequences which are co-expressed
CC   with matrix-remodeling genes. Matrix-remodeling is associated with the
CC   construction, destruction and reorganisation of extracellular matrix
CC   components. The matrix-remodeling-associated nucleotides of the invention
CC   are useful for screening for and purifying ligands that specifically bind
CC   to the nucleotides of the invention. The matrix-remodeling-associated
CC   nucleotides of the invention are also useful in the diagnosis, prognosis,
CC   prevention, treatment and evaluation of therapies for diseases associated
CC   with matrix remodelling (e.g. angiogenesis, arthritis, atherosclerosis,
CC   cancer, cardiomyopathy, diabetic necrosis, fibrosis and ulceration). The
CC   present DNA sequence represents a human matrix-remodeling-associated
CC   nucleotide of the invention.
XX
DR   WPI; 2002-338319/37.
DR   P-PSDB; AAO14785.
XX
SQ   Sequence 1439 BP; 392 A; 336 C; 338 G; 373 T; 0 other;
```

```
Query Match         100.0%;  Score 1439;  DB 24;  Length 1439;
Best Local Similarity   100.0%;
Matches 1439;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy          1 GAGTATCCCTGTTTAATCACTTTTGTGGTTAAAAGAGAGACAGGCCATGAA 50
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGTATCCCTGTTTAATCACTTTTGTGGTTAAAAGAGAGACAGGCCATGAA 50
```

```
Qy         61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTTGTTTCTCCTGTCTCAAGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCCTTGAAGAGTTTAGCCCTGGCTCACTTTTGTTTCTCCTGTCTCAAGAA 120
```

```
Qy        121 AGAAGAAAAAAAGAGACAAAATTACCCAGAAAACCCTCCCTTCCCCACATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAAGAAAAAAAGAGACAAAATTACCCAGAAAACCCTCCCTTCCCCACATG 180
```

```
Qy        181 CAAAATGTTAATTTCCTAGAAAAATCCTTCAGAACCTGAACGACTATGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAAAATGTTAATTTCCTAGAAAAATCCTTCAGAACCTGAACGACTATGGTC 240
```

```
Qy        241 CTCAGGGGTGGCTTCTGCGGTCCCCGCCCAGGCCCCAGACCTGAAGCCCCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCAGGGGTGGCTTCTGCGGTCCCCGCCCAGGCCCCAGACCTGAAGCCCCT 300
```

```
Qy        301 GTTCCTCCCCGGGGACTTCCAGAATTTCTCCTCCAAAAGGAAAGAAAACAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTCCTCCCCGGGGACTTCCAGAATTTCTCCTCCAAAAGGAAAGAAAACAG 360
```

```
Qy        361 GTTGGCAAAACCGCAGGGGCCCGGCCTCCCAAAAAACCCCATGTGTGTACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTGGCAAAACCGCAGGGGCCCGGCCTCCCAAAAAACCCCATGTGTGTACG 420
```

```
Qy        421 GTCCCCAGGCCTTCCCAGCCGCAAACTTAAAGACAGCAGGGCCTTTGCTGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCCCAGGCCTTCCCAGCCGCAAACTTAAAGACAGCAGGGCCTTTGCTGA 480
```

```
Qy        481 GCCAGCTGGGCTTTTTTAAACAACCTAGAGAGACTTTCCCGGAGCTGCCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAGCTGGGCTTTTTTAAACAACCTAGAGAGACTTTCCCGGAGCTGCCTG 540
```

```
Qy        541 GGAAAACGGGGCTTGCCAGAGACACTCACAGTTTCCTTCATGGGTCCCCTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAAAACGGGGCTTGCCAGAGACACTCACAGTTTCCTTCATGGGTCCCCTA 600
```

```
Qy        601 GAATCTCCACATCATTGTCTTTCTTCTGTGCCTTTTCCTTGGTGACCAGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAATCTCCACATCATTGTCTTTCTTCTGTGCCTTTTCCTTGGTGACCAGAA 660
```

```
Qy        661 GTTCCAAGCCTCTAAAAAATGTGCTTTGTGGATCAGAGAGTGCCGCTCCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTCCAAGCCTCTAAAAAATGTGCTTTGTGGATCAGAGAGTGCCGCTCCAA 720
```

```
Qy        721 GCTGCCCTTTCGAGCCCAGTGAGCTCAGCCTCCCAGGCTTTAAAGCCACAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTGCCCTTTCGAGCCCAGTGAGCTCAGCCTCCCAGGCTTTAAAGCCACAT 780
```

```
Qy        841 GAAAAGTAGGTTTCTTGGCCTTGAGTGAGACTGGCGTTCATTAAAGAAGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAAAAGTAGGTTTCTTGGCCTTGAGTGAGACTGGCGTTCATTAAAGAAGAG 900
```

```
Qy        781 AGAAAGGCGTCGAGGAGCTCCGCAGGTTCATTAAAGAAGGCAAAGCACTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGAAAGGCGTCGAGGAGCTCCGCAGGTTCATTAAAGAAGGCAAAGCACTGG 840
```

```
Qy        901 GTACGTAAAACAAAATAGGGCTTGGCCTGGCTGCCTTGATTTTAGTGAGGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTACGTAAAACAAAATAGGGCTTGGCCTGGCTGCCTTGATTTTAGTGAGGG 960
```

```
Qy        961 TGGATGGATGGAATGTATGATGAATAGATAGATAGTGTTGTTTGCATGGTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGGATGGATGGAATGTATGATGAATAGATAGATAGTGTTGTTTGCATGGTA 1020
```

```
Qy       1021 ACAAAACCAAAGCTGATTGGAACAATTAATTGTTGGGGTGTTCTGAGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAAAACCAAAGCTGATTGGAACAATTAATTGTTGGGGTGTTCTGAGGGGG 1080
```

```
Qy       1081 TTTGGGCAGCTTTGAGAGCGGTACAAGAGTTCTTGTGCCTGTGGTCCAGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTTGGGCAGCTTTGAGAGCGGTACAAGAGTTCTTGTGCCTGTGGTCCAGCC 1140
```

```
Qy       1141 AGCCAGTGCATTATTTTAAGTCTTAGAAGCAACTCCTTGGCCCCAGGAATG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGCCAGTGCATTATTTTAAGTCTTAGAAGCAACTCCTTGGCCCCAGGAATG 1200
```

```
Qy       1201 CTGAGATGGGGTCCACGCATCTTCTTCCTTCCTTCTCTCGGGGTACTCGTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGAGATGGGGTCCACGCATCTTCTTCCTTCCTTCTCTCGGGGTACTCGTG 1260
```

```
Qy       1261 CCTCTGCCGCCCATTCTCTTCTCCACGCATCATCCATGAGCTTTAATTCAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCTCTGCCGCCCATTCTCTTCTCCACGCATCATCCATGAGCTTTAATTCAC 1320
```

```
Qy       1321 GGTACGTTCCATAAAGCCAGTATTACACTTAAATGAAGTATTCTTTTTTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGTACGTTCCATAAAGCCAGTATTACACTTAAATGAAGTATTCTTTTTTTT 1380
```

```
Qy       1381 TTTAGAAGGTAAACAAATTTAATAAGCTACCAATATAATGAGAAAAAAAAA 1439
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTTAGAAGGTAAACAAATTTAATAAGCTACCAATATAATGAGAAAAAAAAA 1439
```

RESULT 3
ABX77619

```
ID   ABX77619 standard; cDNA; 1439 BP.
XX
AC   ABX77619;
XX
DT   09-APR-2003  (first entry)
XX
DE   Differentially expressed breast cancer associated cDNA #114.
XX
KW   Breast cancer; differential gene expression; BC-cDNA;
KW   breast cancer diagnosis; breast cancer monitoring;
KW   breast cancer treatment; breast cancer staging; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002156263-A1.
XX
PD   24-OCT-2002.
XX
PF   04-OCT-2001; 2001US-0974298.
XX
PR   05-OCT-2000; 2000US-238331P.
XX
PA   (CHEN/) CHEN H.
XX
PI   Chen H;
XX
DR   WPI; 2003-182653/18.
XX
PT   New cDNAs, which are differentially expressed in (metastatic) breast
PT   cancer useful for diagnosing or staging, breast cancer, or for
PT   monitoring the treatment of breast cancer in an individual -
XX
PS   Claim 1; SEQ ID NO 148; 30pp; English.
XX
CC   The invention describes a combination of cDNAs (designated BC-cDNAs),
CC   which are differentially expressed in breast cancer. The combination
CC   includes 152 cDNA sequences, or their complements. The protein encoded
CC   by any of these BC-cDNAs is useful for screening several molecules or
CC   compounds to identify at least one ligand that specifically binds the
CC   protein, producing or preparing polyclonal or monoclonal antibodies, or
CC   purifying antibodies from a sample. The antibodies, which specifically
CC   bind the protein differentially expressed in breast cancer is useful for
CC   detecting the expression of a protein in a sample. The BC-cDNAs are
CC   also useful for diagnosing, monitoring the treatment of, or staging,
CC   breast cancer. This sequence represents a differentially expressed
CC   breast cancer associated cDNA.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=20020156263.
XX
SQ   Sequence 1439 BP; 392 A; 336 C; 338 G; 373 T; 0 other;
```

```
Query Match             100.0%;  Score 1439;  DB 25;  Length 1439;
Best Local Similarity   100.0%;
Matches 1439;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 GAGTATCCCCTGTTTAATCACTTTTGTGGTTAAAGAGAGACCTTCAGTCTGGCCTCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGTATCCCCTGTTTAATCACTTTTGTGGTTAAAGAGAGACCTTCAGTCTGGCCTCA 60

Qy         61 TTCCTTGAAGAGTTTAGCCCGGCCTCAACTTTTCACTCTATTTCTTCTGTCTCAAGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCCTTGAAGAGTTTAGCCCGGCCTCAACTTTTCACTCTATTTCTTCTGTCTCAAGAA 120

Qy        121 AGAAGAAAAAAGAGACAAAATTACCCAGAAAACCCTCCCCACATGGAGGCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAAGAAAAAAGAGACAAAATTACCCAGAAAACCCTCCCCACATGGAGGCCTTGG 180

Qy        181 CAAATGTTAATTTTCCTAGAAAATCCTTCAGACCTGAAGACCCAGAAAAGAAATCTGGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAAATGTTAATTTTCCTAGAAAATCCTTCAGACCTGAAGACCCAGAAAAGAAATCTGGCT 240

Qy        241 CTCAGGGGTGGGCTTTTGTTGGCGTCCCCGCGCCCGGCGCCCCAGACTAGGTCACGGCGGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCAGGGGTGGGCTTTTGTTGGCGTCCCCGCGCCCGGCGCCCCAGACTAGGTCACGGCGGG 300

Qy        301 GTTCCTCTCCCCGGGGGACTCCAGAATTTCTCTCCTCAAGGAAAGAAAAACAGGGCATGGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTCCTCTCCCCGGGGGACTCCAGAATTTCTCTCCTCAAGGAAAGAAAAACAGGGCATGGC 360

Qy        361 GTTGGCAAAACGCAGGGCCGGCTTCCCAAAAACCCATGTGGTACGATTAAAAGTTGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTGGCAAAACGCAGGGCCGGCTTCCCAAAAACCCATGTGGTACGATTAAAAGTTGGCC 420

Qy        421 GTCCCCAGGGGCTTCCCAGAGCCAAACTTAAAGAGACAGGGGCTTTGCTGAAAACCAAACATGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCCCCAGGGGCTTCCCAGAGCCAAACTTAAAGAGACAGGGGCTTTGCTGAAAACCAAACATGG 480

Qy        481 GCCAGCTGGGCTTTTAACAACCTAGAGACTTTCCGGAGCTGCCTCGAACAGAGCCTCGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAGCTGGGCTTTTAACAACCTAGAGACTTTCCGGAGCTGCCTCGAACAGAGCCTCGG 540

Qy        541 GGAAACAGGGGCGTTGGCCAGAGACAACTCACAGTTCCTTCATGGCCTGTTTGGTCCCCTAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAAACAGGGGCGTTGGCCAGAGACAACTCACAGTTCCTTCATGGCCTGTTTGGTCCCCTAA 600

Qy        601 GAATCCTCCACATCATTGTGCTTTTCCTTGTGCCTTTTCCTTGGTGAGCAACAGAAAGGGAAGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAATCCTCCACATCATTGTGCTTTTCCTTGTGCCTTTTCCTTGGTGAGCAACAGAAAGGGAAGG 660

Qy        661 GTTCCAAGCCTCTAAAAAATGTGCTTTTGTGATCAGGAGTGCGGTTCAAACCAAATACGCGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTCCAAGCCTCTAAAAAATGTGCTTTTGTGATCAGGAGTGCGGTTCAAACCAAATACGCGC 720

Qy        721 GCTGCCCTTTCGAGGCAGTGAGCTTCAGCCTCCAAGGGCTTTAAAGCCACATTTCAGCAAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTGCCCTTTCGAGGCAGTGAGCTTCAGCCTCCAAGGGCTTTAAAGCCACATTTCAGCAAG 780

Qy        781 AGAAAGCGCTGAGAGCTCGCGACAGGTTCATTAAAGAAAGGCCAAAGCCGTTTCTCTCTTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGAAAGCGCTGAGAGCTCGCGACAGGTTCATTAAAGAAAGGCCAAAGCCGTTTCTCTCTTA 840

Qy        841 GAAAAGTAGGTTTTCTTCGGCTTGATGTAGACTGGCTTGCTTGATTTTTAGTGGAAGGGAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAAAAGTAGGTTTTCTTCGGCTTGATGTAGACTGGCTTGCTTGATTTTTAGTGGAAGGGAAT 900

Qy        901 GTACCTAAAACAAAAATAGGGCTTGGCTGGCCTCAAAGGACAAGACAAGGGATGGATGGA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTACCTAAAACAAAAATAGGGCTTGGCTGGCCTCAAAGGACAAGACAAGGGATGGATGGA 960

Qy        961 TGGGATGGATGGATGTATGGATGAATAGATAGATAGTAGTAGTTGCATGTAAATTGCAGAGAAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGGGATGGATGGATGTATGGATGAATAGATAGATAGTAGTAGTTGCATGTAAATTGCAGAGAAA 1020

Qy       1021 ACAAAAACCAAAGCTGATTGGAAACAAAATTAATTAGTGTGGTGTCTGAGGGGGAAGTCGCAGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAAAAACCAAAGCTGATTGGAAACAAAATTAATTAGTGTGGTGTCTGAGGGGGAAGTCGCAGC 1080

Qy       1081 TTTGGGCAGCTTTGAGAAGCGGTACAGAGTTCGTGTGCCGTGTGTCAGCCCTGGAGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTTGGGCAGCTTTGAGAAGCGGTACAGAGTTCGTGTGCCGTGTGTCAGCCCTGGAGCC 1140

Qy       1141 AGCCAGTGCATTATTTAATTTAAGCTCTTAGAAGCAAACTCCTTGGCCGCCCAGGAAATGCGTGACCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGCCAGTGCATTATTTAATTTAAGCTCTTAGAAGCAAACTCCTTGGCCGCCCAGGAAATGCGTGACCC 1200

Qy       1201 CTGAGATGGGTCCAGCGACATCTCTTACACTCCTTCTTCTCCGTGGGGATACTGGACTCGTG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGAGATGGGTCCAGCGACATCTCTTACACTCCTTCTTCTCCGTGGGGATACTGGACTCGTG 1260

Qy       1261 CCTCTGCGGCCCATTCTCTTCTTCAGGCATATCCAATGAGCTTTAATTCACTTTCTGATCAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCTCTGCGGCCCATTCTCTTCTTCAGGCATATCCAATGAGCTTTAATTCACTTTCTGATCAC 1320

Qy       1321 GGTACGTCCATAAAGCCAGTATTACACTTAAAATGAAGTATTCTTTTTGTAATCGTTTTT 1380
```

OM nucleic - nucleic search, using sw model

Run on: August 13, 2003, 04:30:47 ; Search time 5268.5 Seconds
(without alignments)
11173.755 Million cell updates/sec

GenCore version 5.1.6

Title: US-09-818-143-6
Perfect score: 1439
Sequence: 1 gagtatccctcgtttaatca.........agaaaaaaaaaaaaaa 1439

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045813386 residues

1 number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :
        GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pri:*
        10: gb_ro:*
        11: gb_sts:*
        12: gb_sy:*
        13: gb_un:*
        14: gb_vi:*
        15: gb_vr:*
        16: em_ba:*
        17: em_fun:*
        18: em_hum:*
        19: em_in:*
        20: em_mu:*
        21: em_om:*
        22: em_or:*
        23: em_ov:*
        24: em_pat:*
        25: em_ph:*
        26: em_pl:*
        27: em_ro:*
        28: em_sy:*
        29: em_un:*
        30: em_vi:*
        31: em_vi:4
        32: em_htg_hum:*
        33: em_htg_inv:*
        34: em_htg_mus:*
        35: em_htg_other:*
        36: em_htg_pln:*
        37: em_htg_rod:*
        38: em_htg_mam:*
        39: em_htg_vrt:*
        40: em_sy:*
        41: em_htgo_hum:*
        42: em_htgo_mus:*
        43: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1314.8	91.4	1557	9	BC030030	BC030030 Homo sapi
2		1311.4	91.1	2017	9	BC049201	BC049201 Homo sapi
3		1283.6	89.2	128618	9	AC011485	AC011485 Homo sapi
4		1264	87.8	4772	6	AR220892	AR220892 Sequence
5		1263.6	87.8	5370	6	AB040951	AB040951 Homo sapi
6		1259.2	87.5	1943	9	HSM801004	HSM801004 Homo sapi
7	c	706.2	49.1	219026	2	AC144655	AC144655 Callicebu
8		629.8	43.8	633	6	AX440306	AX440306 Sequence
9		370.4	25.7	397	6	AX440306	AX440306 Sequence
10		300.4	20.9	309	6	BD058661	BD058661 Secreted
11		300.4	20.9	309	6	AX330093	AX330093 Sequence
12	c	165.8	11.5	620	11	AX334468	AX334468 Sequence
13	c	165.8	11.5	89328	9	CPA411544	AJ411544 Canis fam
14		115.4	8.0	89328	9	AJ411544	AJ411544 Canis fam
15		115.4	8.0	215105	2	AX329823	AX329823 Sequence
16		86.2	6.0	220731	2	Z93784	Z93784 Human DNA s
17	c	78	5.4	325808	5	AC073717	AC073717 Mus muscu
18	c	78	5.4	203058	5	AC097166	AC097166 Rattus no
19	c	78	5.4	197200	2	AC134186	AC134186 Rattus no
20	c	61.2	4.3	7218	2	AC119556	AC119556 Rattus no
21	c	61.2	4.2	186752	9	AL844150	AL844150 Zebrafish
22	c	60.6	4.2	112270	9	BX005260	BX005260 Danio rer
23		60.4	4.1	179844	9	I66494	I66494 Sequence 14
24	c	58.4	4.1	187465	2	AC019274	AC019274 Homo sapi
25	c	58.4	4.1	204623	2	AC093332	AC093332 Homo sapi
26	c	58.4	4.1	125020	2	AP002501	AP002501 Homo sapi
27	c	57.6	4.0	254127	2	AC090209	AC090209 Homo sapi
28		57	4.0	203817	9	AC132482	AC132482 Mus muscu
29	c	56.8	3.9	63459	9	BX248239	BX248239 Danio rer
30	c	56.8	3.9	81078	2	AC021210	AC021210 Homo sapi
31	c	56.6	3.9	185531	2	AC023588	AC023588 Homo sapi
32	c	56.6	3.9	199030	2	BX470204	BX470204 Danio rer
33	c	56.6	3.9	258571	2	AC111324	AC111324 Rattus no
34	c	56.6	3.9	103019	2	AC097144	AC097144 Rattus no
35		56.4	3.9	154262	2	AC091292	AC091292 Takifugu
36	c	56.4	3.9	172174	5	AC127629	AC127629 Rattus no
37	c	56.4	3.9	210483	2	BX000472	BX000472 Zebrafish
38		56	3.9	155361	2	AC111508	AC111508 Rattus no
39	c	55.8	3.9	119721	2	AC142174	AC142174 Rattus no
40	c	55.8	3.9	168893	2	AC093655	AC093655 Homo sapi
41	c	55.8	3.9	208360	9	BX530090	BX530090 Mus muscu
42	c	55.8	3.9	208360	9	AC020708	AC020708 Homo sapi
43	c	55.8	3.9	218868	2	AL929073	AL929073 Mus muscu
44	c	55.8	3.9	234288	2	AL772311	AL772311 Mus muscu
45	c	55.8	3.9	246337	2	AC106464	AC106464 Rattus no

ALIGNMENTS

RESULT 1
BC030030
LOCUS   BC030030   1557 bp   mRNA   linear   PRI 20-MAY-2002
DEFINITION Homo sapiens, DKFZp434N161 protein, clone MGC:3507 IMAGE:3605290,
            mRNA, complete cds.
ACCESSION   BC030030
VERSION   BC030030.1 GI:20987193
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1557)
AUTHORS   Strausberg,R.
TITLE   Direct Submission

JOURNAL　　　Submitted (06-MAY-2002) National Institutes of Health, Mammalian
　　　　　　　　Gene Collection (MGC), Cancer Genomics Office, National Cancer
　　　　　　　　Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
　　　　　　　　USA
REMARK　　　　NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT　　　Contact: MGC help desk
　　　　　　　　Email: cgapbs-r@mail.nih.gov
　　　　　　　　Tissue Procurement: ATCC
　　　　　　　　cDNA Library Preparation: Rubin Laboratory
　　　　　　　　cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
　　　　　　　　DNA Sequencing by: Institute for Systems Biology
　　　　　　　　http://www.systemsbiology.org
　　　　　　　　contact: anadan@systemsbiology.org
　　　　　　　　Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
　　　　　　　　Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

　　　　　　　　Clone distribution: MGC clone distribution information can be found
　　　　　　　　through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
　　　　　　　　Series: IRAL Plate: 11 Row: P Column: 24
　　　　　　　　This clone was selected for full length sequencing because it
　　　　　　　　passed the following selection criteria: matched mRNA gi: 7661583.

FEATURES　　　Location/Qualifiers
source　　　　1..1557
　　　　　　　　/organism="Homo sapiens"
　　　　　　　　/mol_type="mRNA"
　　　　　　　　/db_xref="LocusID:25959"
　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　/clone="MGC:3507 IMAGE:3605290"
　　　　　　　　/tissue_type="uterus, endometrium adenocarcinoma"
　　　　　　　　/clone_lib="NIH_MGC_44"
　　　　　　　　/lab_host="DH10B-R"
　　　　　　　　/note="Vector: pOTB7"
CDS　　　　　　275..574
　　　　　　　　/codon_start=1
　　　　　　　　/product="DKFZP434N161 protein"
　　　　　　　　/protein_id="AAH30030.1"
　　　　　　　　/db_xref="GI:20987194"
　　　　　　　　/translation="MEALANVNFPRKSFRPEDAGKESGSGQGGFCVPAARFQTMVTGPS
　　　　　　　　CSSPGLQNFSPQRKENRACACWQNAGPAPKNPMCVRLKVGRPQASQRKLKETGLC"
BASE COUNT　　441 a　　367 c　　381 g　　368 t
ORIGIN

Query Match　　91.4%;　　Score 1314.8;　　DB 9;　　Length 1557;
Best Local Similarity　99.7%;　Pred. No. 0;
Matches 1328; Conservative　0; Mismatches　2; Indels　2; Gaps　1;

Ov　　　108　CCTGTCTCAAGAAGAGAAGAAGAAAAAAAGAGAACAAAATACCCAGAAAACCCTCCCTTCCCCAC　167
Db　　　215　CCTGTCTCAAGAAGAAGAAGAAAAAAAGAGAACAAAATACCCAGAAAACCCTCCCTTCCCCAC　274

Ov　　　168　ATGGAGGAGGCCTTGGCAAATGTTAATTTCCTAGAAAATCCTAGAGAACGGAGGA　227
Db　　　275　ATGGAGGAGGCCTTGGCAAATGTTAATTTCCTAGAAAATCCTAGAGAACGGAGGA　334

Qy　　　228　AAAGAAATCTGGCTCTCAGGGGTGGGCTTCTGCGTCCCGGGCCAGGCGCCCAGACTATGTC　287
Db　　　335　AAAGAAATCTGGCTCTCAGGGGTGGGCTTCTGCGTCCCGGGCCAGGCGCCCAGACTATGGTC　394

Qy　　　288　ACAGGGCCGGTCCTGGTTCCTCCCGGGACTCCAGAATTTCTCTCTCAAAGGAAAGAAAAC　347
Db　　　395　ACAGGGCCGGTCCTGGTTCCTCCCGGGACTCCAGAATTTCTCTCTCAAAGGAAAGAAAAC　454

Qy　　　348　AGGGGCATGGCGCTTGTTGGGCAAAACGCAGGGCGGGCCGGCTCCCAAAAAACCCCATGTGTGTACGA　407
Db　　　455　AGGGCATGGCGCTTGTTGGGCAAAACGCAGGGCGGGCCGGCTCCCAAAAAACCCCATGTGTGTACGA　514

Qy　　　408　TTAAAAGTTGGCGGTCCCCAGGCCTTCCCAGGCGGCGAAACTTAAAGACAGGGCTTTGGCTGA　467
Db　　　515　TTAAAAGTTGGCGGTCCCCAGGCCTTCCCAGGCGGCGAAACTTAAAGACAGGGCTTTGGCTGA　574

Qy　　　468　AAACCAAACATGGGCCCAGCTGGGCCTTTTAACAACCTAGAGACTTTCCGGAGCTGCCTGG　527
Db　　　575　AAACCAAACATGGGCCCAGCTGGGCCTTTTAACAACCTAGAGACTTTCCGGAGCTGCCTGG　634

Qy　　　528　AACAGAGCCTGCGCGGGAAACGGGGGCGCTTGCCGAGGAGCACTCACAGTTTCTTCATTGGCCTGT　587
Db　　　635　AACAGAGCCTGCGCGGGAAACGGGGGCTTGCCGAGGAGCACT--CAGTTTCTTCATTGGCCTGT　692

Qy　　　588　TTTGGTTCCCCTAAGAATCTCCACACATCATTGTTCTTTCTTCTTTCTTCCTTGTGGGTGGAGCGA　647
Db　　　693　TTTGGTTCCCCTAAGAATCTCCACATCATTGTTCTTTCTTCTTTCTTCCTTGTGGGTGGAGCGA　752

Qy　　　648　CAGAAAGGGAAGGGTTCCAAGCCTCTAAAAAATGTGCTTTGTTGATCAAGAGAGTGCGGTCCCAA　707
Db　　　753　CAGAAAGGGAAGGGTTCCAAGCCTCTAAAAAATGTGCTTTGTTGATCAAGAGAGTGCGGTCCCAA　812

Qy　　　708　ACCAAAATACGGCGCGGCTGCCTTTCTTTCGAGCGCCAGTGAGCTCAGCCTCTCCAAGGCTTTAAAGCC　767
Db　　　813　ACCAAAATACGGCGCGGCTGCCTTTCTTTCGAGCGCCAGTGAGCTCAGCCTCTCCAAGGCTTTAAAGCC　872

Qy　　　768　ACATTTCAGCAAGAGAAAAGCCGCTGAGAGCTCGCAGGTTCATTAAAGAAGGCAAAGACACTG　827
Db　　　873　ACATTTCAGCAAGAGAAAAGCCGCTGAGAGCTCGCAGGTTCATTAAAGAAGGCAAAGACACTG　932

Qy　　　828　GTTTCTCTCCCTTAGAAAAAGTAGGTTTCTTGGCTTGATGTCAGAGTGAGTGACTGGGCTTTGCTTTGATTTT　887
Db　　　933　GTTTCTCTCCCTTAGAAAAAGTAGGTTTCTTGGCTTGATGTCAGAGTGAGTGACTGGGCTTTGCTTTGATTTT　992

Qy　　　888　TAGTGACAGGGAATAGTACGTAAAAACAAAATAGGGCTTGGCGTGGTCAAAGGGAGAGACAAGCAGG　947
Db　　　993　TAGTGACAGGGAATAGTACGTAAAAACAAAATAGGGCTTGGCGTGGTCAAAGGGAGAGACAAGCAGG　1052

Qy　　　948　ATGGATGGATGGATGGATGATGATGATGATGATAGATAGATAGATAGATAGATGTGTTGCATGTA　1007
Db　　　1053　ATGGATGGATGGATGGATGATGATGATGATGATAGATAGATAGATAGATAGATGTGTTGCATGTA　1112

Qy　　　1008　AATTGCACAGAGAAAACAAAAACCAAAGCTGATTGGAACAAATTAATTGTGGGGTGTCTGAAGGG　1067
Db　　　1113　AATTGCACAGAGAAAACAAAAACCAAAGCTGATTGGAACAAATTAATTGTGGGGTGTCTGAAGGG　1172

Qy　　　1068　GGAAGGGTCGCAGCTTTGGGCCAGCTTTGAGAGAGCGGTACAAGAGTTCTGTTGCCTGTTGTGTC　1127
Db　　　1173　GGAAGGGTCGCAGCTTTGGGCCAGCTTTGAGAGAGCGGTACAAGAGTTCTGTTGCCTGTTGTGTC　1232

Qy　　　1128　CAGCCCCTGGAGCCAGCCAGTGCATTTATTTAAGCTCCTAGAAGGCAACCTCCTTGGCCCAG　1187
Db　　　1233　CAGCCCCTGGAGCCAGCCAGTGCATTTATTTAAGCTCCTAGAAGGCAACCTCCTTGGCCCAG　1292

Qy　　　1188　GAATGGCTGACCCCTGAGATGGGGTCCACGCACAGCTCCTTCTTCACACTTCCTTCTCTCCGTGGGA　1247
Db　　　1293　GAATGGCTGACCCCTGAGATGGGGTCCACGCACAGCTCCTTCTTCACACTTCCTTCTCTCCGTGGGA　1352

Qy　　　1248　TACTGGGACTCGTGCCTCTCGCGCCCATTCTCTTCTTCTCACCGCATATCCCATGAGCTTTAATTTC　1307
Db　　　1353　TACTGGGACTCGTGCCTCTCGCGCCCATTCTCTTCTTCTCACCGCATATCCCATGAGCTTTAATTTC　1412

Qy　　　1308　ACTTTCTGATCACGGTACGTCATCAATAAAGCACAGTATTACACTTAAAATGAAGTATTCTTTTTT　1367
Db　　　1413　ACTTTCTGATCACGGTACGTCATCAATAAAGCACAGTATTACACTTAAAATGAAGTATTCTTTTTT　1472

Qy　　　1368　TGTAATCGTTTTTTTTAGAAAGGTAAACAAAATTAATAAAGCTACCAAAATAATGAGAAAAAA　1427
Db　　　1473　TGTAATCGTTTTTTTTAGAAAGGTAAACAAAATTAATAAAGCTACCAAAATAATGAGAAAAAA　1532

Qy　　　1428　AAAAAAAAAAAA　1439
Db　　　1533　AAAAAAAAAAAA　1544